DOCUMENT RESUME

ED 107 216 IR 001 952

AUTHOR Bower, Cathleen
TITLE Analysis of Complex Intervention Effects in Time-Series Experiments.
PUB DATE Apr 75
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30 through April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *Correlation; *Data Analysis; Economics; Education; *Goodness of Fit; Intervention; Matrices; Psychology; Research Design; Research Methodology; Research Tools; *Statistical Analysis; Statistical Data; *Trend Analysis
IDENTIFIERS ARIMA; Autoregressive Integrated Moving Averages; Least Squares Regression; Regression Analysis; *Time Series Data

ABSTRACT

An iterative least squares procedure for analyzing the effect of various kinds of intervention in time-series data is described. There are numerous applications of this design in economics, education, and psychology, although until recently, no appropriate analysis techniques had been developed to deal with the model adequately. This paper presents and develops a complex example of time-series experiments using simulated data with the intent of illustrating the analytic power of the technique for educational methodologists in a number of substantive fields. The simulated data were developed to conform to an autoregressive integrated moving averages (ARIMA) model, and three intervention effects were built into the series. The first intervention exerted a constant effect; the second damped the effect, and the third caused a general drift (trend shift) of the data points. The methodology and results of the simulation--including the simulation data--are provided. (Author/DGC)

ED107216.

ANALYSIS OF COMPLEX INTERVENTION EFFECTS IN TIME-SERIES EXPERIMENTS

Cathleen Bower

Western Interstate Commission for Higher Education

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

IR 001952

Introduction

An iterative least squares procedure for analyzing the effect of various kinds of intervention in time-series data was originally developed by Box and Tiao (1965) and has recently been extended by Glass, Willson, and Gottman (1974). This procedure provides a method of analysis appropriate for the time-series quasi-experimental design of Campbell and Stanley (1963) where the data consist of a series of observations with one or more interventions (treatment) in the series. There are numerous applications of this design in economics, education, and psychology, although until recently, no appropriate analysis techniques had been developed to deal with the model adequately.

This paper presents and develops a complex example of a time-series experiment using simulated data with the intent of illustrating the analytic power of a time-series experimental analysis for educational methodologists in a variety of substantive fields.

The simulated data were developed to conform to an autoregressive integrated moving averages (ARIMA) model of order (2,1,1); i.e., the autoregressive process was of second order, and the differencing and moving averages processes were of first order. Three intervention effects were built into the 200 observations after the 50th, 100th, and 150th data points. The first intervention exerted a constant effect on the series; the second, a damped effect, and the third caused a general drift (trend shift) of the data points.

While it is true that in practice one would rarely encounter three such divergent intervention effects on the same series of data, these three effects were selected for illustration, and do, in fact, represent commonly encountered intervention or treatment effects in real data. A constant intervention effect can be illustrated by the hypothetical situation where in an all-male college, the level (mean) of SAT scores for freshmen increases by a constant amount (measured over a number of years) due to the intervention of allowing the university to become co-educational. An intervention that exerts a damped effect on the initial level of a series is illustrated by the well-known phenomenon of the Hawthorne effect in behavioral research. In this situation, the scores on tests in a particular classroom might initially rise after introduction of a new classroom approach (treatment or intervention), but would gradually decline to the original level over time as the effect of the intervention or the novelty of the situation wore off. The third type of intervention illustrated in this paper, that of a trend or drift effect, is found in practical situations in which the treatment not only changes the initial mean of the series, but continues to exert an effect by changing the slope of the observations. For example, the distance a student in a physical education class can run might be constant in the first several weeks of the class. Introduction of a required period of calisthenics preceding each attempt to run might enable the student to increase his running speed by a few seconds every week for the remainder of the semester.

Method

A computer program was developed to generate data conforming to the ARIMA (2,1,1) model described above. The two autoregressive parameters were set at -.3 and .6, respectively, and the moving averages parameter at .5; the initial level of the series was set to zero and the error variance to unity. The three interventions, after the 50th, 100th, and 150th observations, consisted of:

(1) a constant effect of 10 points; i.e.,

$$z_t^* = z_t + 10 \qquad t = 51 \text{ to } 100$$

(2) a damped effect; i.e.,

$$\left.\begin{aligned} z^*_{2K+99} &= z_{2K+99} + 20(\tfrac{1}{2})^{2K-1} \\ z^*_{2K+100} &= z_{2K+100} + 20(\tfrac{1}{2})^{2K-1} \end{aligned}\right\} \quad K = 1, 25$$

(3) a trend effect; i.e.,

$$z_t^* = z_t + \sum_{i=1}^{t} (.2) \qquad t = 151 \text{ to } 200$$

Autocorrelations and partial autocorrelations were then computed for each set of 50 observations separately. Table 1 shows these results.

Identification of the proper model from the pattern of autocorrelation and partial autocorrelation coefficients becomes increasingly difficult as the true model increases in complexity. In the present example, the model is quite complex and consequently exact model identification is difficult. Fortunately, there are frequently two or more models of the ARIMA (p,d,q) form that will satisfactorily fit the data. To illustrate this point, three different models were identified from the pattern of autocorrelation and

Table 1a

Autocorrelation and Partial Autocorrelation Coefficients

for Observations 1 to 50

(N=50)

Autocorrelations and Their Standard Errors

Data	Lag																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
z	-.28	.76	-.43	.52	-.53	.44	-.45	.42	-.33	.39	-.34	.28	-.39	.26	-.36	.25	-.32	.25	-.30	.28
σ_r	.14	.15	.22	.23	.25	.28	.29	.30	.31	.32	.33	.34	.34	.34	.35	.36	.37	.37	.37	.38
∇z	-.91	.88	-.84	.79	-.80	.73	-.70	.66	-.59	.58	-.54	.51	-.54	.52	-.50	.49	-.47	.47	-.46	.46
σ_r	.14	.23	.29	.34	.38	.41	.43	.46	.48	.49	.50	.52	.53	.54	.55	.56	.57	.57	.58	.59
$\nabla^2 z$	-.97	.92	-.88	.84	-.81	.77	-.73	.68	-.63	.60	-.57	.55	-.55	.54	-.52	.51	-.49	.48	-.48	.47
σ_r	.14	.24	.31	.36	.40	.43	.46	.48	.50	.52	.53	.54	.56	.57	.58	.59	.60	.60	.61	.62

Partial Autocorrelations and Their Standard Errors

Data	$\hat{\phi}_{11}$	$\hat{\phi}_{22}$	$\hat{\phi}_{33}$	$\hat{\phi}_{44}$
z	-.28	.74	-.30	-.06
∇z	-.91	.28	.00	-.01
$\nabla^2 z$	-.97	-.32	.00	.00

$\sigma_{\hat{\phi}jj} = .15$

Table 1b

Autocorrelation and Partial Autocorrelation Coefficients

for Observations 5' to 100

(N=50)

Autocorrelations and Their Standard Errors

Data	Lag																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
z	-.05	.80	-.07	.59	-.07	.37	-.09	.24	-.15	.10	-.19	.04	-.21	.03	-.25	.01	-.29	-.04	-.35	-.10
σ_r	.14	.14	.21	.21	.24	.24	.26	.26	.26	.26	.26	.27	.27	.27	.27	.27	.27	.28	.28	.29
∇z	-.91	.82	-.72	.63	-.52	.43	-.38	.35	-.31	.28	-.27	.24	-.24	.25	.26	.27	-.26	.28	-.26	.23
σ_r	.14	.23	.29	.32	.34	.36	.37	.38	.38	.39	.39	.40	.40	.40	.41	.41	.41	.42	.42	.42
$\nabla^2 z$	-.95	.85	-.74	.64	-.53	.44	-.39	.35	-.32	.29	-.27	.25	-.25	.25	-.27	.27	-.28	.28	-.27	.24
σ_r	.14	.24	.30	.33	.36	.37	.38	.39	.40	.40	.41	.41	.42	.42	.42	.42	.43	.43	.44	.44

Partial Autocorrelations and Their Standard Errors

Data	$\hat{\phi}_{11}$	$\hat{\phi}_{22}$	$\hat{\phi}_{33}$	$\hat{\phi}_{44}$
z	-.05	-.80	-.07	-.05
∇z	-.91	-.05	.00	-.02
$\nabla^2 z$	-.95	-.52	.00	-.02

$\sigma_{\hat{\phi}_{jj}} = .15$

Table 1c

Autocorrelation and Partial Autocorrelation Coefficients

for Observations 101 to 150

(N=50)

Autocorrelations and Their Standard Errors

Data	Lag																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
z	.69	.77	.51	.56	.39	.40	.28	.27	.20	.22	.21	.21	.14	.10	.01	-.01	-.07	-.09	-.19	-.19
σ_r	.14	.20	.25	.27	.29	.30	.31	.32	.32	.33	.33	.33	.33	.34	.34	.34	.34	.34	.34	.34
∇z	-.61	.69	-.48	.40	-.27	.21	-.15	.06	-.10	-.01	-.01	.05	-.03	.09	-.14	.16	-.21	.24	-.24	.20
σ_r	.14	.19	.23	.25	.27	.27	.28	.28	.28	.28	.28	.28	.28	.28	.28	.28	.28	.29	.29	.29
$\nabla^2 z$	-.89	.76	-.60	.45	-.31	.22	-.14	.07	-.04	.00	.01	.01	-.05	.11	-.18	.24	-.29	.31	-.29	.26
σ_r	.14	.23	.28	.30	.32	.32	.33	.33	.33	.33	.33	.33	.33	.33	.33	.33	.34	.34	.35	.35

Partial Autocorrelations and Their Standard Errors

Data	$\hat{\phi}_{11}$	$\hat{\phi}_{22}$	$\hat{\phi}_{33}$	$\hat{\phi}_{44}$
z	.69	.55	-.06	.01
∇z	-.61	.51	.03	-.16
$\nabla^2 z$	-.89	-.15	.01	-.02

$\sigma_{\hat{\phi}_{jj}} = .15$

Table 1d

Autocorrelation and Partial Autocorrelation Coefficients

for Observations 151 to 200

(N=50)

Autocorrelations and Their Standard Errors

Data	Lag																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
z	.85	.85	.75	.71	.64	.60	.54	.50	.43	.40	.34	.30	.22	.17	.10	.05	.00	-.04	-.07	-.12
σ_r	.14	.22	.28	.32	.35	.37	.39	.40	.42	.42	.43	.44	.44	.44	.44	.44	.44	.44	.44	.45
∇z	-.79	.67	-.52	.35	-.23	.10	.00	.00	.09	-.11	.08	-.07	.01	.04	.03	.01	-.01	-.03	.04	-.07
σ_r	.14	.21	.25	.27	.28	.29	.29	.29	.29	.29	.29	.29	.29	.29	.29	.29	.29	.29	.29	.29
$\nabla^2 z$	-.88	.71	-.55	.39	-.24	.10	-.02	-.05	.11	-.13	.11	-.06	.01	.00	.03	-.02	.02	-.05	.08	-.09
σ_r	.14	.23	.27	.29	.30	.31	.31	.31	.31	.31	.31	.31	.31	.31	.31	.31	.31	.31	.31	.31

Partial Autocorrelations and Their Standard Errors

Data	$\hat{\phi}_{11}$	$\hat{\phi}_{22}$	$\hat{\phi}_{33}$	$\hat{\phi}_{44}$
z	.85	.44	-.01	.02
∇z	-.79	.11	.02	.10
$\nabla^2 z$	-.88	-.26	.00	-.07

$\sigma_{\hat{\phi}jj} = .15$

partial autocorrelation coefficients. These were:

(1) The true ARIMA model (2,1,1) where,

$$z_t = L + a_t - \theta_1 a_{t-1} + (\phi_1 + 1)(z_{t-1} - L) + (\phi_2 - \phi_1)(z_{t-2} - L) - \phi_2 (z_{t-3} - L)$$

(2) The ARIMA model (2,0,1) where,

$$z_t = L + a_t - \theta_1 a_{t-1} + \phi_1(z_{t-1} - L) + \phi_2(z_{t-2} - L)$$

(3) The ARIMA model (1,1,1) where,

$$z_t = L + a_t - \theta_1 a_{t-1} + (\phi_1 + 1)(z_{t-1} - L) - \phi_2(z_{t-2} - L)$$

in each model, z_t is the observation at time t

a_t is the error component at time t

L is the initial level

θ_1 is the first order moving averages parameter

ϕ_1 is the first order autoregressive parameter

ϕ_2 is the second order autoregressive parameter

The simulated data were then run through a computer program for analyzing time-series experiments to compute least squares estimates of the initial level of the series and of the three intervention effects. Three such computer runs were made to obtain the parameter estimates for each of the potential models identified from the autocorrelation and partial autocorrelation coefficients. In each case, a design matrix was constructed (as input for the program) which specified the form of the parameter effects to be estimated. (The design matrix is shown in the appendix.)

The final step in this project was to assess the goodness of fit of each of the least squares solutions. To accomplish this, the residuals after fitting the model were computed for each observation. These residual data were then autocorrelated, and a χ^2 statistic computed which tested the independence of the residuals. (The χ^2 statistic is a function of the sum of squared autocorrelation coefficients.)

Results and Discussion

Table 2 summarizes the results of analyzing the data as time-series experiments under three different ARIMA models. All three produced parameter estimates of initial level and intervention effects near the true values built into the data with the exception of the (2,0,1) model estimate of initial level. The true initial level of the series was zero, and the (1,1,1) and (2,1,1) models yielded estimates of .41 and .50, respectively, both not significantly different from zero. The (2,0,1) model, however, estimated the initial level of the series as 2.61 which was significantly different from zero with a t statistic value of 3.96. Undoubtedly, this estimate was large because the original data required differencing in order to achieve stationarity, yet the (2,0,1) model specifies that no differencing is necessary.

The estimated values of the first intervention effect were 9.74, 10.01, and 9.87, respectively, for the (2,1,1), 2,0,1), and (1,1,1) models, all of which were highly significantly different from zero and near the true value of 10.0. Similarly, the estimates for the second intervention effect were all near 10.0 and highly significantly different from zero. The three models

TABLE 2

Estimates of Initial Level and Intervention Effects
With t Statistics for Three ARIMA Models

ARIMA Model	Error Variance	ϕ_1	ϕ_2	θ_1	Initial Level	t	I_1	t_1	I_2	t_2	I_3	t_3	X^2
(2,1,1)	.902	-.6	.3	.1	.41	.53	9.74	13.00	9.86	11.94	.31	3.36	37.10
(2,0,1)	1.027	.0	.9	-.3	2.61	3.96	10.01	13.75	10.00	11.06	.46	10.83	45.94
(1,1,1)	.918	-.9	***	-.1	.50	.67	9.87	14.68	9.96	12.22	.30	3.86	38.19

*** indicates model does not contain this parameter

produced estimates of the third intervention parameter of .31, .46, and .30, respectively, for the (2,1,1), (2,0,1), and (1,1,1) model, all of which were reasonably close to the true value of 0.2.

The error variance estimates were also near the true value of unity in all three cases, the smallest value yielded by the (2,1,1) model, the middle value by the (1,1,1), and the largest value by the (2,0,1) model.

The χ^2 test statistics of the residuals after fitting each model were all non-significant indicating that any of the three models was adequate for describing the data according to the χ^2 criterion; the (2,1,1) model had the lowest χ^2 value, however.

The estimates of ϕ_1, ϕ_2, and θ_1 for each model were, of course, quite different across the three models. Probably the most interesting of these, is the result for the (2,1,1) model, since the true autoregressive and moving averages parameters were specified in the construction of the model, and can be compared to the estimated values. The values were originally specified as $\phi_1 = -.3$, $\phi_2 = .6$, and $\theta_1 = .5$. The minimum error variance was found at $\phi_1 = -.6$, $\phi_2 = .3$, and $\theta_2 = .1$. The differences between the two sets of values probably resulted from sampling error during generation of the data (using the random number generator to build random error into the model).

Summary

Simulated data were generated to conform to an ARIMA (2,1,1) process, and three intervention effects were built into the series. Analysis of the series was performed in two stages. First, separate correlograms of the data were inspected before and after each intervention to verify that the time-series conformed to an ARIMA model of order (2,1,1).

Second, the data were analyzed using a program specifically developed for analyzing time-series data with one or more interventions. A design matrix was specified for the four parameters to be estimated (i.e., the initial level of the series, and the three intervention effects), incorporating a constant, a damped, and a change in drift intervention effect. Values of the two autoregressive parameters, ϕ_1, and ϕ_2, and the moving averages parameter, θ_1, were examined over the invertibility and stationarity regions to determine the minimum residual error term, and thus, the best fit for the model, using the least squares criterion. The values of ϕ and θ which minimized the residual error variance were found to be similar to and within an acceptable range of the values used to build the simulated data. Examination of the corresponding parameter estimates and associated t statistics for the initial level of the series and the three intervention effects also revealed nearly exact estimates of the true parameter values which were highly significant for all design effects.

Appendix: Design Matrix Used for the Three Time-Series Experiment Analyses

1.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00

Appendix: Design Matrix (continued)

1.00	1.00	0.00	0.00
1.00	1.00	0.00	0.00
1.00	1.00	0.00	0.00

Appendix: Design Matrix (continued)

1.00	0.00	.00	0.00
1.00	0.00	.00	0.00
1.00	0.00	.00	0.00
1.00	0.00	.00	0.00
1.00	0.00	.00	0.00
1.00	0.00	.00	0.00
1.00	0.00	-0.00	0.00
1.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00
1.00	0.00

Appendix: Design Matrix (continued)

1.00	0.00	0.00	27.00
1.00	0.00	0.00	28.00
1.00	0.00	0.00	29.00
1.00	0.00	0.00	30.00
1.00	0.00	0.00	31.00
1.00	0.00	0.00	32.00
1.00	0.00	0.00	33.00
1.00	0.00	0.00	34.00
1.00	0.00	0.00	35.00
1.00	0.00	0.00	36.00
1.00	0.00	0.00	37.00
1.00	0.00	0.00	38.00
1.00	0.00	0.00	39.00
1.00	0.00	0.00	40.00
1.00	0.00	0.00	41.00
1.00	0.00	0.00	42.00
1.00	0.00	0.00	43.00
1.00	0.00	0.00	44.00
1.00	0.00	0.00	45.00
1.00	0.00	0.00	46.00
1.00	0.00	0.00	47.00
1.00	0.00	0.00	48.00
1.00	0.00	0.00	49.00
1.00	0.00	0.00	50.00

REFERENCES

Box, G.E.P. and Tiao, C.A. A Change in Level of a Nonstationary Time-Series. Biometrika, 1965, 52; 181-192.

Campbell, D.T. and Stanley, J.C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally and Co., 1963.

Glass, G.V, Willson, V.L., and Gottman, J.M. Design and Analysis of Time-Series Experiments. Boulder, Colorado: Colorado Associated University Press, 1974.